

RAW SEQUENCE LISTING

**The Biotechnology Systems Branch of the Scientific and Technical
Information Center (STIC) no errors detected.**

Application Serial Number: 10/532, 013
Source: PLT
Date Processed by STIC: 04/11/2006

ENTERED

PCT

RAW SEQUENCE LISTING

DATE: 04/11/2006

PATENT APPLICATION: US/10/532,013

TIME: 12:44:18

Input Set : A:\pto.da.txt

Output Set: N:\CRF4\04112006\J532013.raw

```

3 <110> APPLICANT: E.I. du Pont de Nemours and Company
4 Hallahan, David L.
6 <120> TITLE OF INVENTION: CIS-PRENYLTRANSFERASES FROM THE RUBBER-PRODUCING PLANTS
RUSSIAN DANDELION
7 (TARAXACUM KOK-SAGHYZ) AND SUNFLOWER (HELIANTHUS ANNUS)
9 <130> FILE REFERENCE: CL2039
C--> 11 <140> CURRENT APPLICATION NUMBER: US/10/532,013
C--> 11 <141> CURRENT FILING DATE: 2005-04-20
11 <160> NUMBER OF SEQ ID NOS: 45
13 <170> SOFTWARE: PatentIn version 3.1
15 <210> SEQ ID NO: 1
16 <211> LENGTH: 746
17 <212> TYPE: DNA
18 <213> ORGANISM: Taraxacum kok-saghyz
20 <400> SEQUENCE: 1
21 gatcgaaggc tttctgaaag aagtttagtat tataaaaccaa tatggcggtta gagtcttggt 60
23 catcggtgat ctcgataggt tatatgagcc cgtaaggatt gctgctgaga aggccatgga 120
25 agccaccgct aaaaactcaa ccacatatct cctcgtatgt gttgcttaca cttcttccca 180
27 tgaatccca cgtgccatcc acgaagcttg tgaagaaagc atacgggtca tgaacggaaa 240
29 cgggtttttc aatggaagcg gatataccaa cgtgaatcat ggaagtcagg cggatgatcaa 300
31 agtgggtgat cttgataagc atatgtacat gggggtggca ccggatcctg atattttagt 360
33 aaggagctcc ggcgaaccaa ggctgagcaa ctttctgctg tggcagacca ccaactgttt 420
35 gttgtattcc ccgaaagctt tgtggccgga gatgggggtt tggcaggtgg tttggggaat 480
37 cttggagtgt caaaacaatt ataattactt ggagaagaag aagaagcagg cgtaaggatg 540
39 tgttcaaaaa gtaaggtaat ctgtctttta atgagtttgg agtgtgctgt gagcattaat 600
41 gggatttttc ttcccaatat gaactttcaa ttttgggtcg attataatat atgatccata 660
43 tgtatatgaa cgttgtgtga tgcattatac gagcagaaga acgttgtatt ttactaaaa 720
45 aaaaaaaaaa aaaaaaaaaa aaaaaa 746
48 <210> SEQ ID NO: 2
49 <211> LENGTH: 788
50 <212> TYPE: DNA
51 <213> ORGANISM: Taraxacum kok-saghyz
53 <400> SEQUENCE: 2
54 gcccttcgcg gatccagacg ctgcgtttgc tggctttgat gaaaataatc tattccacca 60
56 agttatctct ctctctctct ctctctctct ctctctctct ctctctgtct gtctctcttc 120
58 ctctgtctct ctagtataca attggcaaat aggattaagc cggctcattt gttaaaccaa 180
60 gatgcaagtg aatccaatca ttactacaga tagttcactg aaactagtgg aagaagaaag 240
62 atcaaatggt aggatcggca atttcttagg aggcctaaac gccaccttaa gaaaactcgt 300
64 gtttcgtgtc attgcttctc gcccaatccc agaacacatc gccttcatcc tcgatggaaa 360
66 ccgaagggtt gccaggaaat ggaacctcac agaaggcgcc gccacaaaaa ccggcttcct 420
68 agcactcatg tcgggtcctc aatactgcta cgagatcgga gttaagtacg tcaccatcta 480
70 cgccttcagc ctgcacaatt tcaatcgacg ccctgatgaa gtccagtacg tcatggactt 540
72 gatgcaagac aagatcgaa gctttctgaa agaagttagt attataaacc aatatggcgt 600
74 tagagtcttg ttcacgggtg atctcgatag gttatatgag cccgtaagga ttgctgctga 660

```

RAW SEQUENCE LISTING

DATE: 04/11/2006

PATENT APPLICATION: US/10/532,013

TIME: 12:44:18

Input Set : A:\pto.da.txt

Output Set: N:\CRF4\04112006\J532013.raw

```

76 gaaggccatg gaagccaccg ctaaaaactc aaccacatat ctctcgtat gtgttgctta      720
78 cacttcttcc catgaaatcc cagtgccat ccacgaagct tgtgaagaaa gcatacgggt      780
80 catgaacg                                          788
83 <210> SEQ ID NO: 3
84 <211> LENGTH: 906
85 <212> TYPE: DNA
86 <213> ORGANISM: Taraxacum kok-saghyz
88 <400> SEQUENCE: 3
89 atgcaagtga atccaatcat tactacagat agttcactga aactagtgga agaagaaaga      60
91 tcaaatggta ggatcggcaa tttcttagga ggcttaaacy ccaccttaag aaaactcgtg      120
93 tttcgtgtca ttgcttctcg cccaatccca gaacacatcg ccttcacctc cgatggaaac      180
95 cgaaggttcg ccaggaaatg gaacctcaca gaaggcgccg gccacaaaac cggcttccta      240
97 gcactcatgt cggctctcaa atactgtac gagatcggag ttaagtacgt caccatctac      300
99 gccttcagcc tcgacaattt caatcgacgc cctgatgaag tccagtacgt catggacttg      360
101 atgcaagaca agatcgaagg ctttctgaaa gaagttagta ttataaacca atatggcggtt      420
103 agagtcttgt tcacgcgtga tctcgatagg ttatatgagc ccgtaaggat tgctgctgag      480
105 aaggccatgg aagccaccgc taaaaactca accacatata tctcgtatg tgttgcttac      540
107 acttcttccc atgaaatccc acgtgccatc cacgaagctt gtgaagaaag catacgggtc      600
109 atgaacggaa acgggttttt caatggaagc ggatatacca acgtgaatca tggagtcag      660
111 ggggtgatca aagtgggtga tcttgataag catatgtaca tgggggtggc accggatcct      720
113 gatattttag taaggagctc cggcgaaaca aggctgagca actttctgct gtggcagacc      780
115 accaactgtt tgttgatttc cccgaaagct ttgtggcccg agatgggggt ctggcagggtg      840
117 gtttggggaa tcttgaggtt tcaaaacaat tataattact tggagaagaa gaagaagcag      900
119 gcgtaa                                          906
122 <210> SEQ ID NO: 4
123 <211> LENGTH: 301
124 <212> TYPE: PRT
125 <213> ORGANISM: Taraxacum kok-saghyz
127 <400> SEQUENCE: 4
129 Met Gln Val Asn Pro Ile Ile Thr Thr Asp Ser Ser Leu Lys Leu Val
130 1                    5                    10                    15
133 Glu Glu Glu Arg Ser Asn Gly Arg Ile Gly Asn Phe Leu Gly Gly Leu
134                20                    25                    30
137 Asn Ala Thr Leu Arg Lys Leu Val Phe Arg Val Ile Ala Ser Arg Pro
138                35                    40                    45
141 Ile Pro Glu His Ile Ala Phe Ile Leu Asp Gly Asn Arg Arg Phe Ala
142                50                    55                    60
145 Arg Lys Trp Asn Leu Thr Glu Gly Ala Gly His Lys Thr Gly Phe Leu
146 65                70                75                80
149 Ala Leu Met Ser Val Leu Lys Tyr Cys Tyr Glu Ile Gly Val Lys Tyr
150                85                90                95
153 Val Thr Ile Tyr Ala Phe Ser Leu Asp Asn Phe Asn Arg Arg Pro Asp
154                100               105               110
157 Glu Val Gln Tyr Val Met Asp Leu Met Gln Asp Lys Ile Glu Gly Phe
158                115               120               125
161 Leu Lys Glu Val Ser Ile Ile Asn Gln Tyr Gly Val Arg Val Leu Phe
162                130               135               140
165 Ile Gly Asp Leu Asp Arg Leu Tyr Glu Pro Val Arg Ile Ala Ala Glu
166 145                150                155                160

```

RAW SEQUENCE LISTING

DATE: 04/11/2006

PATENT APPLICATION: US/10/532,013

TIME: 12:44:18

Input Set : A:\pto.da.txt

Output Set: N:\CRF4\04112006\J532013.raw

```

169 Lys Ala Met Glu Ala Thr Ala Lys Asn Ser Thr Thr Tyr Leu Leu Val
170                               165                               170                               175
173 Cys Val Ala Tyr Thr Ser Ser His Glu Ile Pro Arg Ala Ile His Glu
174                               180                               185                               190
177 Ala Cys Glu Glu Ser Ile Arg Val Met Asn Gly Asn Gly Phe Phe Asn
178                               195                               200                               205
181 Gly Ser Gly Tyr Thr Asn Val Asn His Gly Ser Gln Ala Val Ile Lys
182                               210                               215                               220
185 Val Val Asp Leu Asp Lys His Met Tyr Met Gly Val Ala Pro Asp Pro
186 225                               230                               235                               240
189 Asp Ile Leu Val Arg Ser Ser Gly Glu Thr Arg Leu Ser Asn Phe Leu
190                               245                               250                               255
193 Leu Trp Gln Thr Thr Asn Cys Leu Leu Tyr Ser Pro Lys Ala Leu Trp
194                               260                               265                               270
197 Pro Glu Met Gly Phe Trp Gln Val Val Trp Gly Ile Leu Glu Phe Gln
198                               275                               280                               285
201 Asn Asn Tyr Asn Tyr Leu Glu Lys Lys Lys Lys Gln Ala
202 290                               295                               300
205 <210> SEQ ID NO: 5
206 <211> LENGTH: 505
207 <212> TYPE: DNA
208 <213> ORGANISM: Helianthus annuus
210 <400> SEQUENCE: 5
211 accagaaaga ttgaagggtt tatgaaagaa ttaacaattg tgaacaggta tgggtgtaga 60
213 gtcttggttta tcggcgatct taaaagggtta tacgagcccg ttagagttgc agccgagaaa 120
215 gcaatggagg ccactgctaa caacacacat acatatcttt tagtatgtgt tgcttacact 180
217 tcttcacacg aaatcccgcg tgccgtttat gaatcttgcg aagaaaagag tggtggaacc 240
219 ggagttatga ttaatggaaa tggaaagtgtg aacggagatt acagtgaaga aaagagtggg 300
221 ggaaccggag ttatggtgaa tggaaatggg agtgtgaatg gagattacag taatggagat 360
223 catgaggagg ggggttaaagt ggtggatatt gacaaacata tgtatatggc agtggctcct 420
225 gatcctgata ttttggtcag gagctcaggg gagacgaggt tgagtaactt tttgctgtgg 480
227 caaaccacca actgcgtggt gtatt 505
230 <210> SEQ ID NO: 6
231 <211> LENGTH: 168
232 <212> TYPE: PRT
233 <213> ORGANISM: Helianthus annuus
235 <400> SEQUENCE: 6
237 Thr Arg Lys Ile Glu Gly Phe Met Lys Glu Leu Thr Ile Val Asn Arg
238 1 5 10 15
241 Tyr Gly Val Arg Val Leu Phe Ile Gly Asp Leu Lys Arg Leu Tyr Glu
242 20 25 30
245 Pro Val Arg Val Ala Ala Glu Lys Ala Met Glu Ala Thr Ala Asn Asn
246 35 40 45
249 Thr His Thr Tyr Leu Leu Val Cys Val Ala Tyr Thr Ser Ser His Glu
250 50 55 60
253 Ile Pro Arg Ala Val Tyr Glu Ser Cys Glu Glu Lys Ser Gly Gly Thr
254 65 70 75 80
257 Gly Val Met Ile Asn Gly Asn Gly Ser Val Asn Gly Asp Tyr Ser Glu
258 85 90 95

```

RAW SEQUENCE LISTING

DATE: 04/11/2006

PATENT APPLICATION: US/10/532,013

TIME: 12:44:18

Input Set : A:\pto.da.txt

Output Set: N:\CRF4\04112006\J532013.raw

```

261 Glu Lys Ser Gly Gly Thr Gly Val Met Val Asn Gly Asn Gly Ser Val
262          100          105          110
265 Asn Gly Asp Tyr Ser Asn Gly Asp His Glu Glu Gly Val Lys Val Val
266          115          120          125
269 Asp Ile Asp Lys His Met Tyr Met Ala Val Ala Pro Asp Pro Asp Ile
270          130          135          140
273 Leu Val Arg Ser Ser Gly Glu Thr Arg Leu Ser Asn Phe Leu Leu Trp
274 145          150          155          160
277 Gln Thr Thr Asn Cys Val Leu Tyr
278          165
281 <210> SEQ ID NO: 7
282 <211> LENGTH: 228
283 <212> TYPE: PRT
284 <213> ORGANISM: Calendula officinalis
286 <400> SEQUENCE: 7
288 Met Pro Lys His Val Ala Phe Ile Met Asp Gly Asn Arg Arg Trp Ala
289 1          5          10          15
292 Val Glu Lys Gly Trp Ser Pro Met Thr Gly His Ser Ala Met Arg Lys
293          20          25          30
296 Thr Leu Gln Ser Leu Leu Phe Arg Cys Ser Lys Phe Lys Ile Lys Ala
297          35          40          45
300 Val Ser Ile Tyr Ala Phe Ser Thr Glu Asn Trp Thr Arg Pro Lys Glu
301          50          55          60
304 Glu Val Asp Phe Leu Met Glu Met Tyr Glu Asp Leu Leu Arg Thr Asp
305 65          70          75          80
308 Ala Glu Glu Leu Leu Ser Leu Gly Cys Arg Val Ser Ile Met Gly Lys
309          85          90          95
312 Lys Thr Asn Leu Pro Lys Ser Leu Gln Lys Leu Cys Ile Glu Ile Glu
313          100          105          110
316 Glu Lys Ser Arg Ala Asn Ser Gly Thr His Val Asn Tyr Ala Leu Asn
317          115          120          125
320 Tyr Ser Gly Lys Tyr Asp Ile Ile Glu Ala Cys Lys Ser Val Ala Thr
321          130          135          140
324 Lys Val Lys Asp Gly Val Ile Ile Pro Lys Gln Ile Asp Glu Lys Tyr
325 145          150          155          160
328 Phe Lys Gln Glu Leu Gly Thr Lys Met Ile Asp Phe Pro Tyr Pro Asp
329          165          170          175
332 Leu Val Ile Arg Thr Ser Gly Glu Ile Arg Leu Ser Asn Phe Met Leu
333          180          185          190
336 Trp Gln Met Ala Tyr Ser Glu Leu Tyr Phe Thr Asp Lys Tyr Phe Pro
337          195          200          205
340 Asp Phe Gly Glu Asn Asp Leu Ile Glu Ala Leu Leu Ala Phe Gln Lys
341          210          215          220
344 Val Arg Lys Cys
345 225
348 <210> SEQ ID NO: 8
349 <211> LENGTH: 290
350 <212> TYPE: PRT
351 <213> ORGANISM: Hevea brasiliensis

```

RAW SEQUENCE LISTING

DATE: 04/11/2006

PATENT APPLICATION: US/10/532,013

TIME: 12:44:18

Input Set : A:\pto.da.txt

Output Set: N:\CRF4\04112006\J532013.raw

353 <400> SEQUENCE: 8

```

355 Met Glu Leu Tyr Asn Gly Glu Arg Pro Ser Val Phe Arg Leu Leu Gly
356 1 5 10 15
359 Lys Tyr Met Arg Lys Gly Leu Tyr Ser Ile Leu Thr Gln Gly Pro Ile
360 20 25 30
363 Pro Thr His Ile Ala Phe Ile Leu Asp Gly Asn Arg Arg Phe Ala Lys
364 35 40 45
367 Lys His Lys Leu Pro Glu Gly Gly Gly His Lys Ala Gly Phe Leu Ala
368 50 55 60
371 Leu Leu Asn Val Leu Thr Tyr Cys Tyr Glu Leu Gly Val Lys Tyr Ala
372 65 70 75 80
375 Thr Ile Tyr Ala Phe Ser Ile Asp Asn Phe Arg Arg Lys Pro His Glu
376 85 90 95
379 Val Gln Tyr Val Met Asp Leu Met Leu Glu Lys Ile Glu Gly Met Ile
380 100 105 110
383 Met Glu Glu Ser Ile Ile Asn Ala Tyr Asp Ile Cys Val Arg Phe Val
384 115 120 125
387 Gly Asn Leu Lys Leu Leu Ser Glu Pro Val Lys Thr Ala Ala Asp Lys
388 130 135 140
391 Ile Met Arg Ala Thr Ala Asn Asn Ser Lys Cys Val Leu Leu Ile Ala
392 145 150 155 160
395 Val Cys Tyr Thr Ser Thr Asp Glu Ile Val His Ala Val Glu Glu Ser
396 165 170 175
399 Ser Glu Leu Asn Ser Asn Glu Val Cys Asn Asn Gln Glu Leu Glu Glu
400 180 185 190
403 Ala Asn Ala Thr Gly Ser Ser Thr Val Ile Gln Thr Glu Asn Met Glu
404 195 200 205
407 Ser Tyr Ser Gly Ile Lys Leu Val Asp Leu Glu Lys Asn Thr Tyr Ile
408 210 215 220
411 Asn Pro Tyr Pro Asp Val Leu Ile Arg Thr Ser Gly Glu Thr Arg Leu
412 225 230 235 240
415 Ser Asn Tyr Leu Leu Trp Gln Thr Thr Asn Cys Ile Leu Tyr Ser Pro
416 245 250 255
419 Tyr Ala Leu Trp Pro Glu Ile Gly Leu Arg His Val Val Trp Ser Val
420 260 265 270
423 Ile Asn Phe Gln Arg His Tyr Ser Tyr Leu Glu Lys His Lys Glu Tyr
424 275 280 285
427 Leu Lys
428 290

```

431 <210> SEQ ID NO: 9

432 <211> LENGTH: 290

433 <212> TYPE: PRT

434 <213> ORGANISM: Hevea brasiliensis

436 <400> SEQUENCE: 9

```

438 Met Glu Leu Tyr Asn Gly Glu Arg Pro Ser Val Phe Arg Leu Leu Glu
439 1 5 10 15
442 Lys Tyr Met Arg Lys Gly Leu Tyr Ser Ile Leu Thr Gln Gly Pro Ile
443 20 25 30
446 Pro Thr His Ile Ala Phe Ile Leu Asp Gly Asn Arg Arg Phe Ala Lys

```

RAW SEQUENCE LISTING ERROR SUMMARY
PATENT APPLICATION: US/10/532,013

DATE: 04/11/2006
TIME: 12:44:19

Input Set : A:\pto.da.txt
Output Set: N:\CRF4\04112006\J532013.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:27; Xaa Pos. 5,9,13,15,16,18,19,20
Seq#:28; Xaa Pos. 2,3,4,7,8,9,11,13,16,17,19,20,21,22,24,35
Seq#:29; Xaa Pos. 2,3,4,5,10,12
Seq#:30; Xaa Pos. 3,4
Seq#:31; Xaa Pos. 1,2,3,7,10,11,13,14,15,18,19,22
Seq#:32; Xaa Pos. 2,3,8,9,10,11,12
Seq#:33; Xaa Pos. 1,3,7,10,12,16,17,20,21,23,25,26,28,29,30,31,34
Seq#:38; Xaa Pos. 4
Seq#:39; Xaa Pos. 2,4,5,6,8,9
Seq#:40; Xaa Pos. 2,5,7,9,12,13,18
Seq#:44; N Pos. 6,8

Invalid Line Length:

The rules require that a line not exceed 72 characters in length. This includes spaces.

Seq#:1; Line(s) 6

VERIFICATION SUMMARY

DATE: 04/11/2006

PATENT APPLICATION: US/10/532,013

TIME: 12:44:19

Input Set : A:\pto.da.txt

Output Set: N:\CRF4\04112006\J532013.raw

L:11 M:270 C: Current Application Number differs, Replaced Current Application No
L:11 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:1824 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:27 after pos.:0
L:1828 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:27 after pos.:16
L:1938 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:28 after pos.:0
L:1942 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:28 after pos.:16
L:1946 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:28 after pos.:32
L:1997 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:29 after pos.:0
L:2024 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:30 after pos.:0
L:2111 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:31 after pos.:0
L:2115 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:31 after pos.:16
L:2172 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:32 after pos.:0
L:2289 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:33 after pos.:0
L:2293 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:33 after pos.:16
L:2297 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:33 after pos.:32
L:2377 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:38 after pos.:0
L:2427 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:39 after pos.:0
L:2483 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:40 after pos.:0
L:2487 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:40 after pos.:16
L:2548 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:44 after pos.:0